Half-time review

Instructor: Jie Wang
H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Tech

Outline

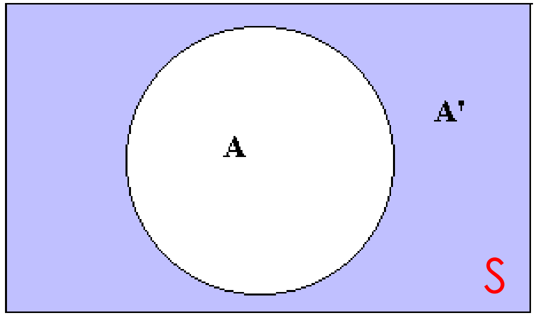
1. Probability Theory
2. Univariate Random Variable
3. Bivariate Random Variable
4. Normal distribution

What is probability theory?

Probability theory is the branch of mathematics concerned with probability, the analysis of random phenomena - wikipedia

1. Random experiment
2. Sample space
3. Event

} Set theory



- Random experiment

    Any procedure that can be repeated infinitely and has more than one possible outcomes.

- Sample space

    The collection of all possible outcomes and is denoted by S in this course.

- Event

    The collection of some possible outcomes in S and is a subset of the sample space S.

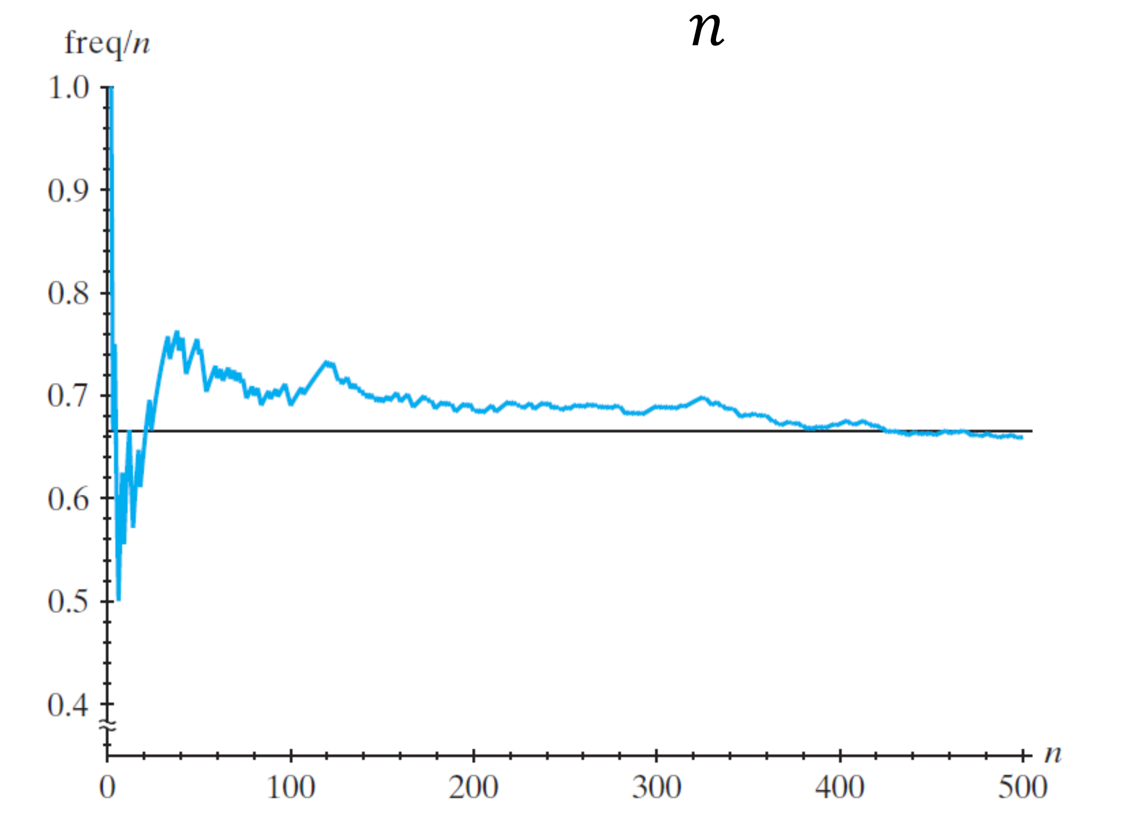
- Event A has happened

    Event A is said to has happened if the outcome of the experiment is in A.

What is probability?

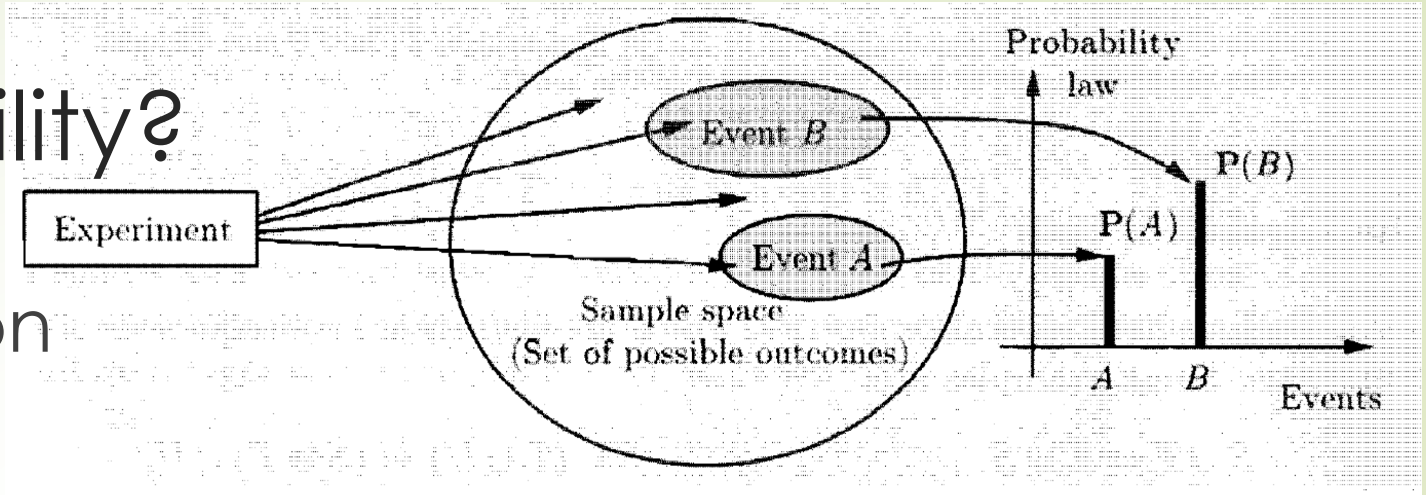
- Relative frequency

$$\lim_{n\to\infty} \frac{number\ of\ times\ that\ event\ A\ has\ happend\ in\ n\ experiments}{n}$$



What is probability?

- Probability function



Definition 1.1-1

Probability is a real-valued set function P that assigns, to each event A in the sample space S, a number $P(A)$, called the probability of the event A, such that the following properties are satisfied:

(a) $P(A) \geq 0$;

(b) $P(S) = 1$;

(c) if $A_1, A_2, A_3, \ldots$ are events and $A_i \cap A_j = \emptyset$, $i \neq j$, then

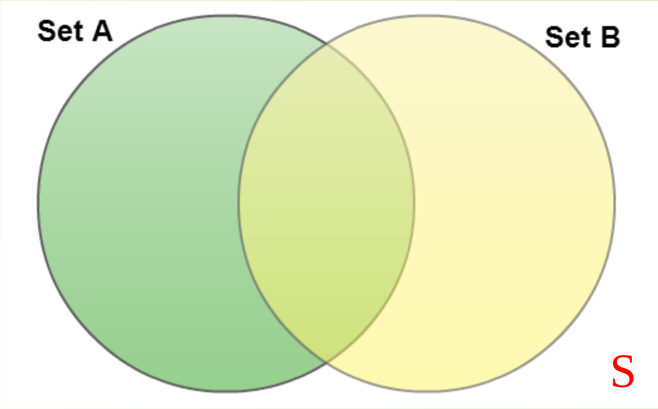
$$P(A_1 \cup A_2 \cup \cdots \cup A_k) = P(A_1) + P(A_2) + \cdots + P(A_k)$$

for each positive integer k, and

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$$

for an infinite, but countable, number of events.

What is probability?



- Properties of Probability function

$P(A) = 1 - P(A')$.

$P(\emptyset) = 0$.

If events A and B are such that $A \subset B$, then $P(A) \leq P(B)$.

For each event A, $P(A) \leq 1$.

If A and B are any two events, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

"Equally likely" and counting techniques

- If the outcomes are "equally likely", i.e.,

$$P(\{e_i\}) = \frac{1}{m}, \qquad i = 1, 2, \ldots, m. \qquad S = \{e_1, e_2, \ldots, e_m\}$$

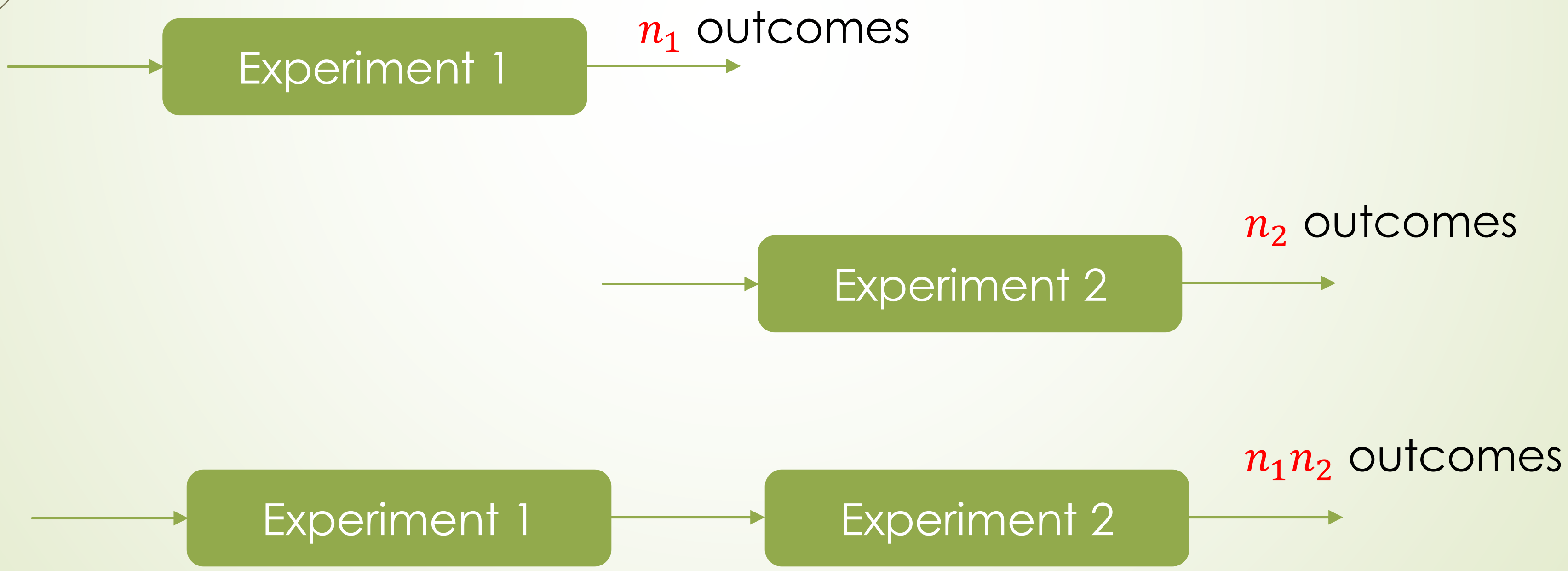
then

$$P(A) = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } S}$$

The problem of computing P(A) becomes the problem of counting the number of outcomes in the set A.

“Equally likely” and counting techniques

- Multiplication principle

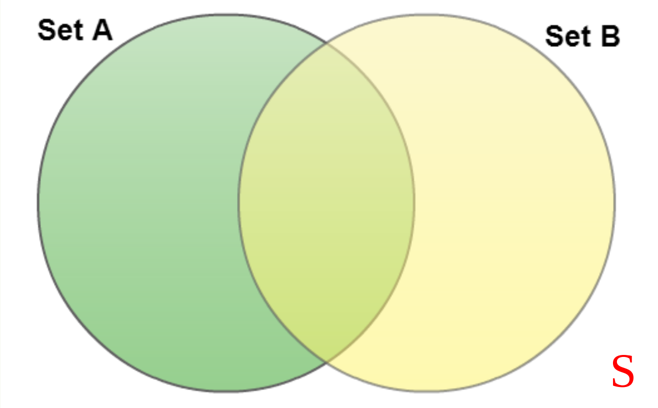


Conditional Probability

- The **conditional probability** of an event *A*, given that event *B* has occurred, is defined by

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

provided that $P(B) > 0$.



1. Conditional probability is a probability.
2. The sample space shrinks from S to B.
3. $P(A \cap B) = P(A)P(B \mid A)$ $P(A \cap B) = P(B)P(A \mid B)$

25 balloons on a board, of which 10 balloons are yellow, 8 are red, and 7 are green. Probability of that the first two balloons hit are yellow?

Independent events

- Events A and B are **independent** if and only if $P(A \cap B) = P(A)P(B)$. Otherwise, A and B are called **dependent** events.

- If A and B are independent events, then the following pairs of events are also independent:

  (a) A and B'; (b) A' and B; (c) A' and B'.

- Events A, B, and C are **mutually independent** if and only if

(a) A, B, and C are pairwise independent; that is,

$$P(A \cap B) = P(A)P(B), \qquad P(A \cap C) = P(A)P(C),$$

and

$$P(B \cap C) = P(B)P(C).$$

(b) $P(A \cap B \cap C) = P(A)P(B)P(C)$.

Bayes Theorem

- Assume that
  1. S is a sample space and B_1, B_2, …, B_m are mutually exclusive and exhaustive w.r.t. the sample space S
  2. The prior probabilities of B_i, i=1,…,m, are positive

Then

a) For any event A

$$P(A) = \sum_{i=1}^{m} P(B_i \cap A)$$
$$= \sum_{i=1}^{m} P(B_i)P(A \mid B_i)$$

Prior probability
Posterior probability
Likelihood

b) If P(A)>0

$$P(B_k \mid A) = \frac{P(B_k)P(A \mid B_k)}{\sum_{i=1}^{m} P(B_i)P(A \mid B_i)}$$

Univariate Random Variable

- pmf and pdf

Given a discrete or continuous RV $X: S \to X(S) \subset R$, or simply X defined on $D \subset R$, we define accordingly a pmf or pdf to assign the probability for the RV:

1. *pmf for discrete RV:* $f(x): D \to [0, 1]$

   $f(x) \geq 0, x \in D; \sum_{x \in \mathrm{D}} f(x) = 1; P(x \in A) = \sum_{x \in A} f(x), A \subset D$

2. *pdf for continuous RV:* $f(x): D \to [0, \infty)$

   $f(x) \geq 0, x \in D; \int_D f(x)dx = 1; P(x \in [a, b]) = \int_a^b f(x)dx$

3. *cdf for RV:* $F(x) = P(X \leq x), x \in D$

   for continuous RV, $F^{(1)}(x) = f(x), x \in D$

Univariate Random Variable

- Mathematical expectation [average value of $u(X)$]

$$E[u(X)] = \begin{cases} \sum_{x \in D} u(x)f(x), & discrete\ RV \\ \int_D u(x)f(x)dx, & continuous\ RV \end{cases}$$

- Properties of mathematical expectation
  1. If c is a constant, $E[c] = c$
  2. If c is a constant and $u(X)$ is a function of X, $E[cu(X)] = cE[u(X)]$
  3. Ifc_1, c_2 are constants and $g_1(X), g_2(X)$ are functions of X,
     $E[c_1 g_1(X) + c_2 g\_2(X)] = c_1 E[g_1(X)]+c_2 E[g_2(X)]$

Characteristics of RV

- Mean [average value of X]

$$E[X]$$

- Variance [measure of the dispersion or spread out of X]

$$Var[X] = E(X - E[X])^2$$

$\sqrt{Var(X)}$ - standard deviation

- rth order Moments

$$E[X^r]$$

- mgf[uniquely characterizes the distribution of the RV]

$M(t) = E(e^{tX})$, $|t| < h$, if there is $h > 0$ such that $E(e^{tX})$ exists.

$$M(0) = 1, M^{(1)}(0) = E[X], M^{(2)}(0) = E[X^2], M^{(r)}(0) = E[X^r]$$

Bivariate Random Variable

- The outcome of the random experiment is a tuple of several things of interests.
- Joint pmf and joint pdf

Given a discrete or continuous RV X defined on D, we define accordingly a pmf or pdf to assign the probability for the RV:

1. *pmf for discrete RV:* $f(x, y): D \rightarrow [0, 1]$

   $f(x, y) \geq 0, (x, y) \in D; \sum_{(x,y) \in D} f(x) = 1,$

   $$P((x, y) \in A) = \sum_{(x,y) \in A} f(x, y), A \subset D$$

2. *pdf for continuous RV:* $f(x, y): D \rightarrow [0, \infty)$

   $f(x, y) \geq 0, (x, y) \in D;\ \iint_D f(x, y) dxdy = 1;$

   $$P((x, y) \in A) = \iint_A f(x, y) dxdy, A \subset D$$

Bivariate Random Variable

- Marginal pmf and marginal pdf

Given two discrete or continuous RVs X, Y defined on D and their joint pmf or joint pdf $f(x, y)$, we define accordingly the marginal pmf or marginal pdf to assign the probability for the RV X:

$$D_X = \{all\ possible\ values\ of\ X\ in\ D\}, D_Y = \{all\ possible\ values\ of\ Y\ in\ D\}$$

1. *marginal pmf for discrete RV:* $f_X(x): D_X \to [0, 1]$

$$f_X(x) = \sum_{y \in D_Y} f(x, y)$$

with the understanding $f(x, y) = 0, (x, y) \notin D$

2. marginal pdf for continuous RV: $f_X(x): D_X \to [0, \infty)$

$$f_X(x) = \int_{D_Y} f(x, y) dy$$

with the understanding $f(x, y) = 0, (x, y) \notin D$

Bivariate Random Variable

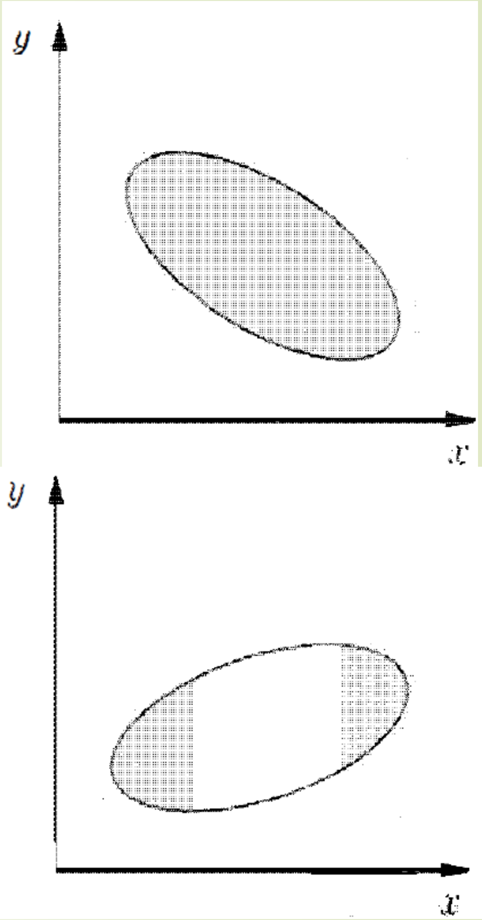
- Mathematical Expectation

$$E[u(X,Y)] = \begin{cases} \sum_{(x,y)\in \mathrm{D}} u(x,y)f(x,y), & \textit{discrete RV} \\ \iint_D u(x,y)f(x,y)dxdy, & \textit{continuous RV} \end{cases}$$

- Covariance and Correlation Coefficient

$$Cov(X,Y) = E[(X - E[x])(Y - E[Y])]$$

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}, \quad |\rho| \le 1$$



1. $Cov(X,Y) > 0, \rho > 0, (X - E[x])$ *and* $(Y - E[Y])$ tend to have the same sign; $\rho = 1 \Longrightarrow X - E[x] = c(Y - E[Y])$ with $c > 0$
2. $Cov(X,Y) < 0, \rho < 0, (X - E[x])$ *and* $(Y - E[Y])$ tend to have the opposite sign; $\rho = -1 \Longrightarrow X - E[x] = c(Y - E[Y])$ with $c < 0$

Bivariate Random Variable

- Independent Random Variables

X and Y are said to be independent if and only if

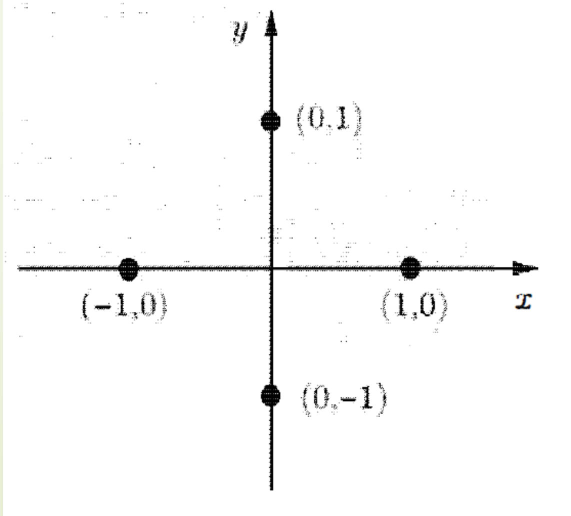
$$f(x, y) = f_X(x) f_Y(y)$$

A necessary condition for X and Y to be independent

$$D = D_X \times D_Y$$

Independence → Uncorrelation

1. The converse is not true in general.
2. The converse is however true for multivariate normal (Gaussian) distribution.



Bivariate Random Variable

- Conditional pmf and Conditional pdf

Given two discrete or continuous RVs X, Y defined on D and their joint pmf or joint pdf $f(x, y)$, and marginal pmf or marginal pdf $f_X(x)$to assign the probability for the RV Y given that $X = x$:

$$h(y|x) = \frac{f(x,y)}{f_X(x)}, f_X(x) > 0$$

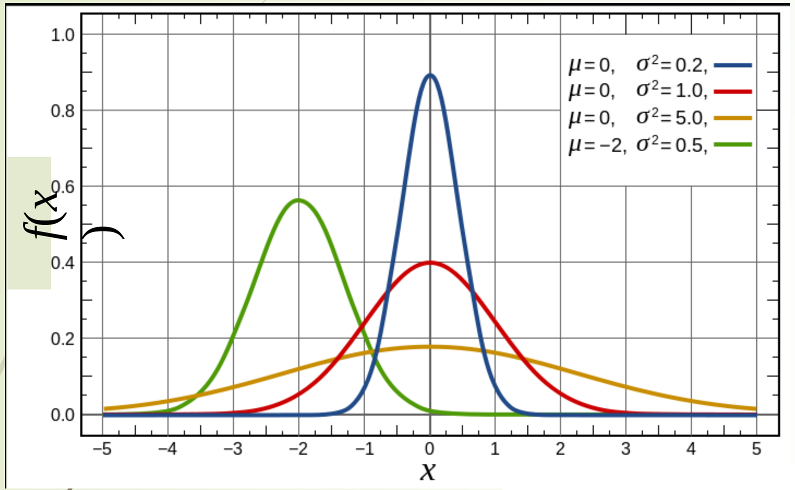
$$g(x|y) = \frac{f(x,y)}{f_Y(y)}, f_Y(y) > 0$$

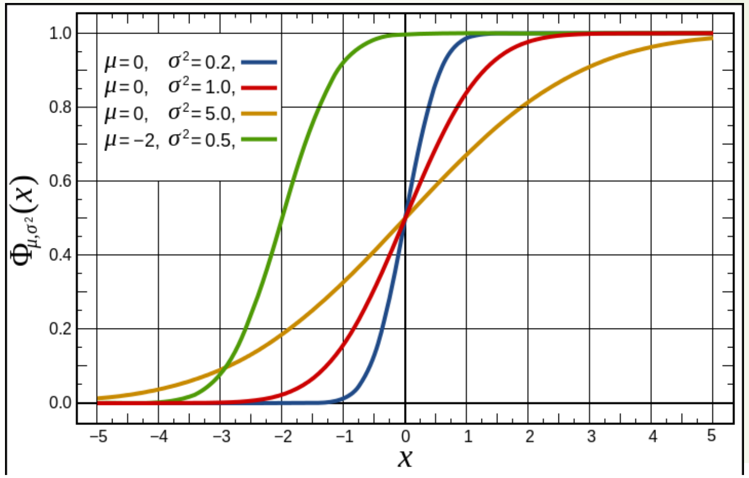
- Conditional mathematical expectation

$$E[u(Y)|X = x] = \begin{cases} \sum_{y \in D_Y} u(y)h(y|x), & discrete\ RV \\ \int_{D_Y} u(y)h(y|x)dx, & continuous\ RV \end{cases}$$

Normal (Gaussian) distribution

- Univariate Normal distribution





$X \sim N(\mu, \sigma^2)$ if $$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \qquad -\infty < x < \infty,$$

$$E[X] = \mu, Var[X] = \sigma^2, M(t) = \exp(\mu t + \frac{1}{2}\sigma^2 t^2)$$

$$\frac{X-\mu}{\sigma} \sim N(0,1) \Longrightarrow$$

$$P(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$P(Z \le z) = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw$$

$$\Phi(-z) = 1 - \Phi(z)$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026

Normal (Gaussian) distribution

- Bivariate Normal distribution
  1. Marginal pdf and conditional pdf are all normal.
  2. Independence is equivalent to uncorrelation!

The End